

WHAT IS CLAIMED IS:

1. A method for identifying among a list of texts those that have edit distance from a search string that is less than a threshold, said method comprising:

beginning with a first text from a list of texts as a computation text;

providing a search string and a threshold value;

performing a column-by-column grid computation of edit distance between the search string and the computation text, stopping early if a column minimum value is at least the threshold;

if the edit distance is computed and is below the threshold, then reporting that the computation text is close to the search string in edit distance;

if all columns are computed, then setting the next computation text to the next text in the list after the present computation text;

if not all columns are computed, then setting the next computation text to the next text in the list after the present computation text that does not share with the present computation text the prefix corresponding to the columns up to and including the column with minimum value at least the threshold;

applying the same procedure as for the first computation text, but re-using columns corresponding to any prefix shared with the present computation text;

continuing until the text list is exhausted.

2. The method of claim 1, further comprising:

ordering the text list to place texts with shared prefixes together in the list.

3. The method of claim 1, wherein the step of performing a column-by-column grid computation of edit distance further comprises the steps of:

after a column is computed, identifying the range of rows from the lowest cell with value less than the limit to the highest cell with value less than the limit;

in the next column, not computing the cells below this range (if the bottom border cell has value at least the threshold); computing normally the cells along this range and one higher, and computing the cells above based only on each cell below, until a cell with at least the limit value is computed.

4. The method of claim 1, further comprising the steps of:

making an alternative text list in which each occurrence of a character in a set of characters is replaced by some character in the set;

when a search is performed, if the search string lacks all characters in said set of characters, then using the alternative text list rather than the original text list.

5. The method of claim 1, wherein the step of performing a column-by-column grid computation of edit distance further comprises the steps of:

if the present computation text is not the first computation text, then re-using a column for the computation text if the previous column is the same as for the previous computation text, and at least one of the following conditions holds:

the characters corresponding to the column in the present computation text and the previous computation text are the same;

the search string lacks the characters corresponding to the column in the present computation text and the previous computation text.